SALT AND WATER MOVEMENT ON HILLSLOPE SOIL TOPOSEQUENCES IN THE BURDEKIN RIVER IRRIGATION AREA

Progress Report - April 1989

by

John Doherty

Report No. 89/02(a)

INTRODUCTION

Since November 1988, work has proceeded on the construction of an automatic calibration computer program for steady-state and transient groundwater flow. This has proceeded simultaneously with other project activities viz. the carrying out of a small electromagnetic survey over part of Leichhardt Downs (see the February 1989 Progress Report) and the development of a steady-state model for Stage 2 of the Leichhardt Downs portion of the BRIA; the latter will be the subject of a later report. The automatic calibration program has now been completed and tested and appears to be error-free. The aim of this report is to provide news of its completion and to relate its capabilities. A full documentation is not provided as this will be covered in a program manual. The latter will be compiled at some later stage when it is felt that no further changes to input/output portions of the program, or changes made to improve efficiency, are required.

THE ALGORITHM

The program provides estimates of aquifer parameters and their stochastic properties through iterative parameter refinement, such that an "objective function" of these parameters is minimized. This function incorporates the (weighted) sum of the squares of the deviations between the measured borehole heads throughout the aquifer and the heads calculated on the basis of the current parameter estimates, as well as a function that varies with the deviations of current parameter estimates from any prior estimates of these parameters, if the latter are provided. As well as allowing the user the ability to incorporate pump-test or any other field-gathered aquifer parameters, the use of prior estimates in the inversion algorithm provides stability in the face of a tendency for oscillatory or non-convergent behaviour due to system non-uniqueness if different parameter types are simultaneously inverted. The governing principles of the estimation process are derived from Maximum Likelihood Theory. The latter is described in detail as it applies in the groundwater modelling context by Carrera and Neuman (1986a, 1986b, 1986c); the present algorithm's capabilities are very similar to those described in these papers, though the forward modelling and optimization routines are quite different.

The forward modelling processor is MODFLOW, a well-known, inexpensive, comprehensively documented and quite general finite difference groundwater package. It has been modified slightly for inclusion in the inversion package in that most "write" statements have been deleted (excluding those pertaining to error conditions), and the program is called as a subroutine by the main program. In order to minimize changes to MODFLOW, data interchange is conducted through reading and writing to files of the type expected by MODFLOW, so that, prior to performing a parameter estimation run, MODFLOW input files are set up in the normal manner (some of these will be changed during the run). An interpolation subroutine is used to obtain model heads at observation bores from the MODFLOW unformatted block head output.

Optimization is performed by adjusting aquifer parameters such that the objective function is minimized. The Gauss-Marquardt method is employed, a method which has proven to be efficient and less prone to instability and non-convergence than other methods, even when parameters are highly correlated; its use is described in Bard (1974). The user is given the option of selecting Marquardt's " " himself at each iteration, or having this done automatically. The step size is computed automatically such that the objective function's minimum in the step direction will be achieved.

The Gauss-Marquardt method requires that derivatives of heads with respect to aquifer parameters at observation points be provided. This is done by varying each parameter, solving for the heads and obtaining derivatives by finite differences; centred differences are used for greater accuracy when convergence has almost been achieved. This method consumes a lot of cpu time when the number of parameters are high. For the confined case, where the flow equations are linear, derivatives could be computed more quickly using a method similar to the Adjoint State method described by Carrera and Neuman. However, this would not have been applicable to unconfined flow so that it was not incorporated in the present program in order to preserve generality.

When optimization of the objective function has been completed, the stochastic properties of the aquifer parameters can be determined. These include the covariance matrix, the parameter correlation matrix, as well as the eigenvectors and eigenvalues of the parameter covariance matrix. From these are obtained indicators of parameter uncertainty and of parameter interdependence; they are only indicators because their derivation assumes a linear relationship between heads and parameters, a property that only approximately exists in the vicinity of the collection of parameter best estimates.

THE PROGRAM CAPABILITIES

At this stage the program's name is GENINV (for GENeral INVerse). GENINV is capable of performing parameter estimation for either steady-state or transient groundwater flow conditions. The program can be run interactively, in which case the Marquardt "λ's" are provided by the user, who can terminate the program when convergence is judged to have occurred, or it can be run in batch mode, in which case these parameters are determined automatically.

Any parameter type can be estimated for which MODFLOW will accept an unformatted spatial distribution array as input. This includes transmissivity, hydraulic conductivity, recharge, storage coefficient, well withdrawal rate, etc. for one or a number of layers. GENINV was written to be as general as possible. Hence parameter types are not named, just the files in which the relevant array of unformatted values are stored; these will correspond to the relevant files nominated in the MODFLOW input. Also required is a matching integer array for each parameter type, the integers representing the spatial discretization of the model area into zones of constant value for each parameter type; these zones do not need to coincide for different parameter types. For each parameter type, individual parameters (corresponding to individual aerial zones) can remain fixed if desired. If they are to be estimated, a prior estimate for that parameter can be given if desired, together with a weighting to be attached to that estimate in the inversion process as well as its correlation with prior estimates that may be available for other parameters of the same type. If a parameter is to be estimated, an initial guess of its value is required.

Parameter values themselves or a transformation of the parameter values can be estimated. Two types of parameter transformation are available, viz. (natural) logarithmic transformation and logistic transformation. In the latter case the transformed parameters (p_t) are related to the untransformed ones (p_o) by

$$p_t = \log [p_o/(1 - p_o)];$$

note that for any p_t, p_o lies in the interval (0,1). For some parameter types it is advantageous to use these transformations as the possibility of their values being estimated as less than zero or greater than 1 is eliminated. Also, using either of these transformations, the relationship between that parameter type and the model heads may more closely approach linearity. As well as this, Carrera and Neuman point out that the field distribution of transmissivity is more likely to be log-normal than normal; hence estimating the log of transmissivity rather than transmissivity itself is more in accord with the theoretical basis of Maximum Likelihood Theory upon which the inversion process is based.

A number of parameters of any type can be grouped such that the relative values of all parameters within that group remain fixed while the multiplying factor applicable to all of them is estimated; again, prior estimates of these factors can be provided and this is taken into account in the inversion process. This grouping can be useful when an estimate of flow into the sides of the model is required. The boundary blocks can be divided into groups and the relative values of inflow into each of the blocks comprising each group can be set according to water level contour shape and spacing. An estimate of the group multiplier (which will yield the total flow into that group of boundary blocks) can then be sought.

For transient flow conditions a correlation matrix between observed heads at individual bores at different times is required. The rationale behind this requirement is that it provides a more realistic description of the likely discrepancies between model and observed heads, and hence is likely to yield more realistic parameter estimates; again, see Carrera and Neuman. The user can opt for no time-correlation simply by entering all non-diagonal elements of the matrix as zero. He may then retain a weighting distribution for measurements at different times if he chooses different values for the non-zero diagonal elements.

AN EXAMPLE

Fig. 1 shows the water level contours observed on 17th July 1984 over Leichhardt Downs, near Ayr, North Queensland. The area covers most of a well-defined basin, with the Burdekin River to the west, Stokes Range to the north and east, and Louisa Mountain to the south. Fig. 2 shows a MODFLOW finite-difference grid covering the area. The shaded blocks are inactive; all cells are 250m x 250m. The active part of the grid is bounded by a series of constant-head cells along the river, some no-flow cells at the north, with lateral inflow cells constituting the remainder of the boundary (the grid boundary coincides with the 50m topographic

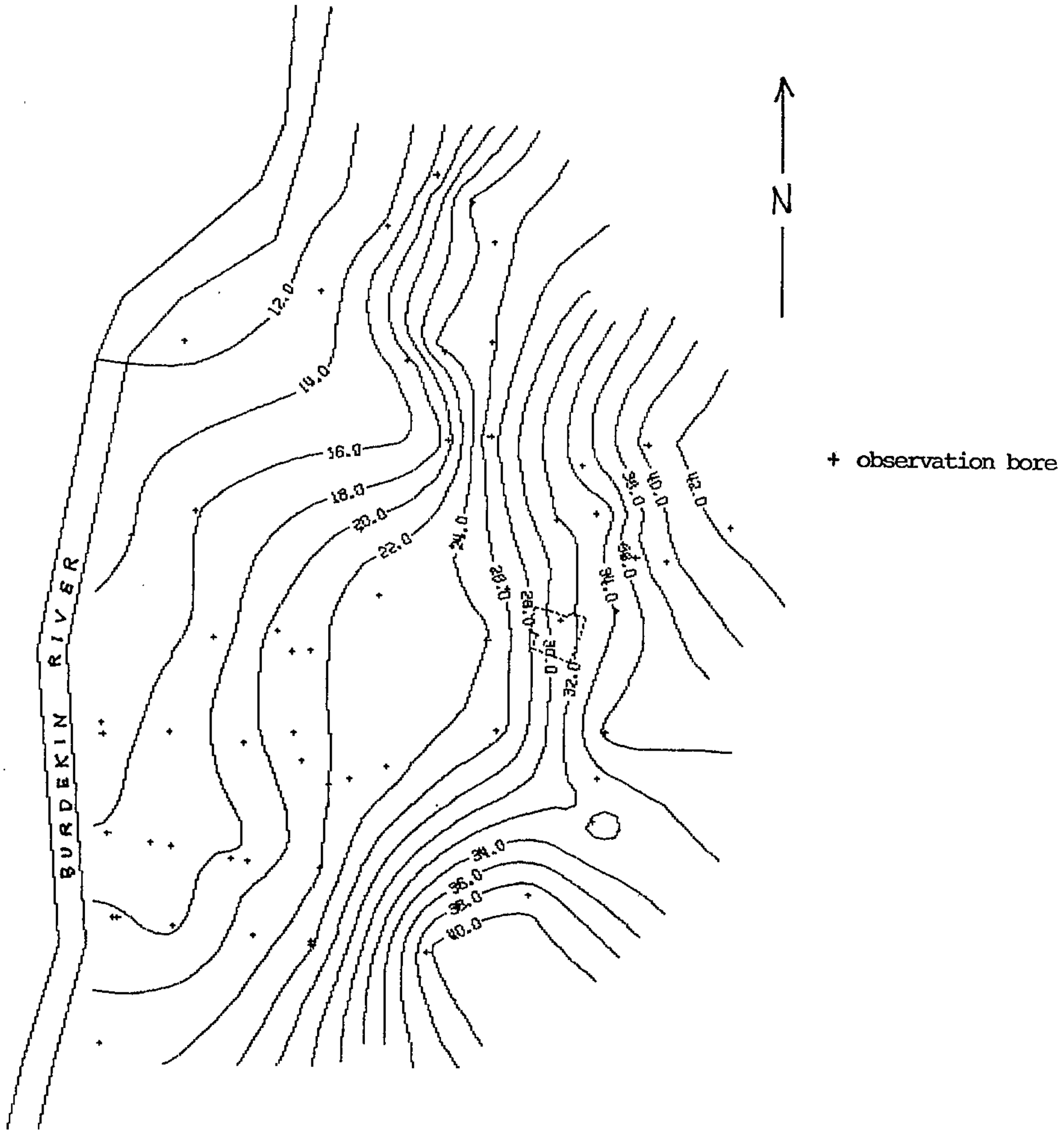

Fig. 1. Water level contours over Leichhardt Downs on 17th July 1984. The scale is 1:100000.

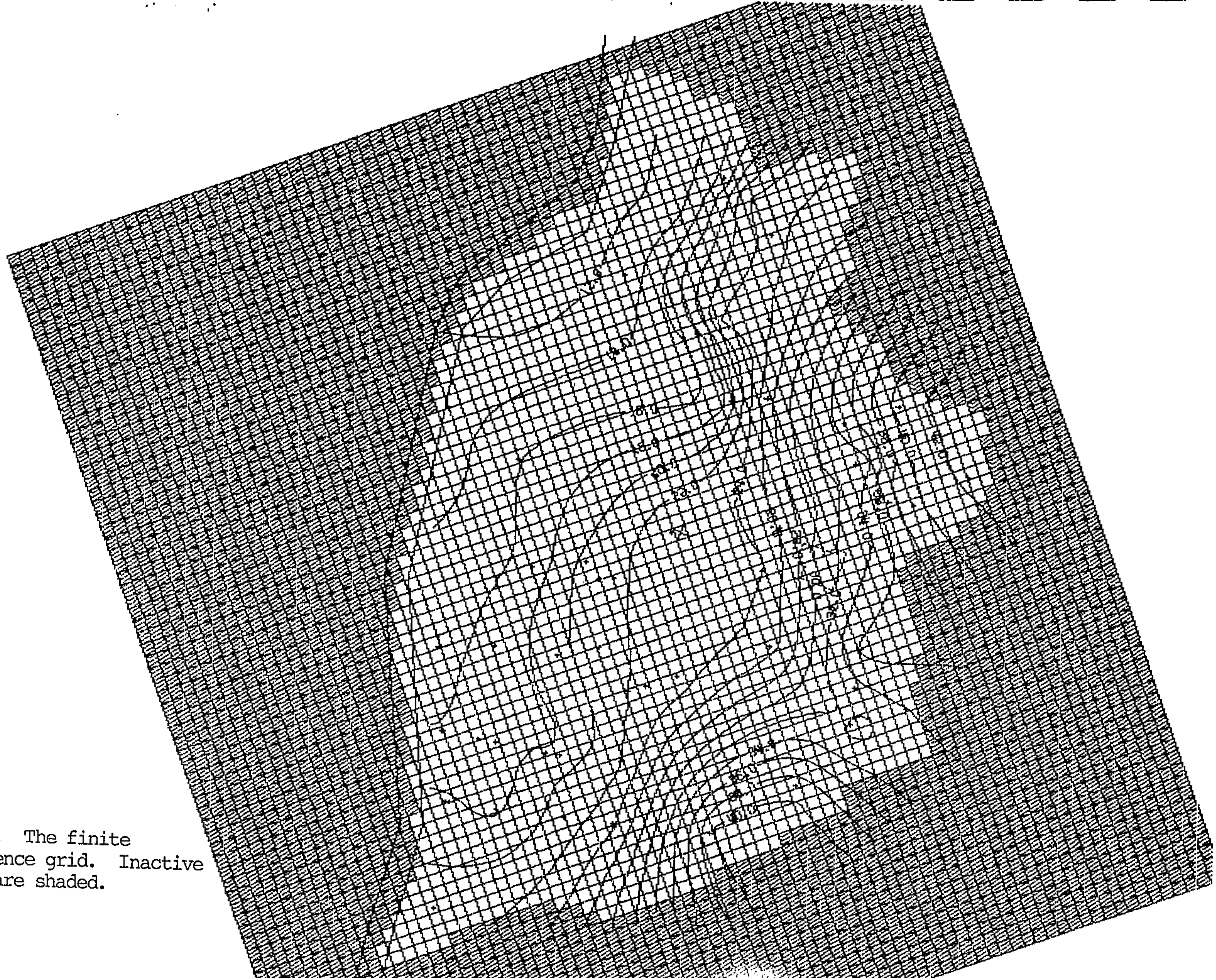

Fig. 2. The finite difference grid. Inactive cells are shaded.

contour). Internal steady-state cell recharge was evaluated on the basis of the Shaw and Thorburn (1985) salt mass-balance model, together with Queensland Department of Primary Industries soil data over the area. The distribution of aquifer transmissivities assuming steady-state flow conditions was sought.

At first, the active portion of the grid was divided into eight zones and GENINV was used to establish the transmissivity of each. However it was soon realized that the area is much more complex than this. The alluvials near the river to the west are underlain by a number of old creek channels, so that alluvial thickness and type is variable. Further east the aquifer is of fractured rock. The country rock is composed of schists and gneisses variably weathered to an average depth of about 20m. These are pervasively intruded by north-south trending andesite dykes. Field observations and drilling have shown that some of these appear to behave as almost impermeable barriers to groundwater flow while others are associated with increases in hydraulic conductivity by virtue of the fracturing within them and in the host rock around them, this fracturing being associated with their emplacement.

After a number of GENINV runs for which the number and location of transmissivity zones was varied, the transmissivity distribution shown in Fig. 3 was achieved; Fig. 4 shows a comparison between field and model contours. For the 64 bores used to generate the contours the sum of the squares of the deviations between field and model heads for the model of Fig. 3 is only 18.7m^2. This is an average deviation of 0.5m per hole; as a total head drop of 30m is represented by the contours over the basin, this is an average error of 1.8% per hole. Such errors could never have been achieved with manual calibration.

Table 1 shows the transmissivities and standard deviations assigned to each zone of Fig. 3. Most of the transmissivity values are realistic, though some are definitely excessive. The latter indicate either errors in head measurements, in assumed lateral or local recharge near the transmissivity zone, erroneous location of zone boundaries (chosen by the user), or inapplicability of steady-state conditions. Far from being a weakness of the automatic calibration algorithm, they indicate where factors external to it are in error; the parameters it chose are those that provide the best possible fit between model and field data (no prior estimates were given).

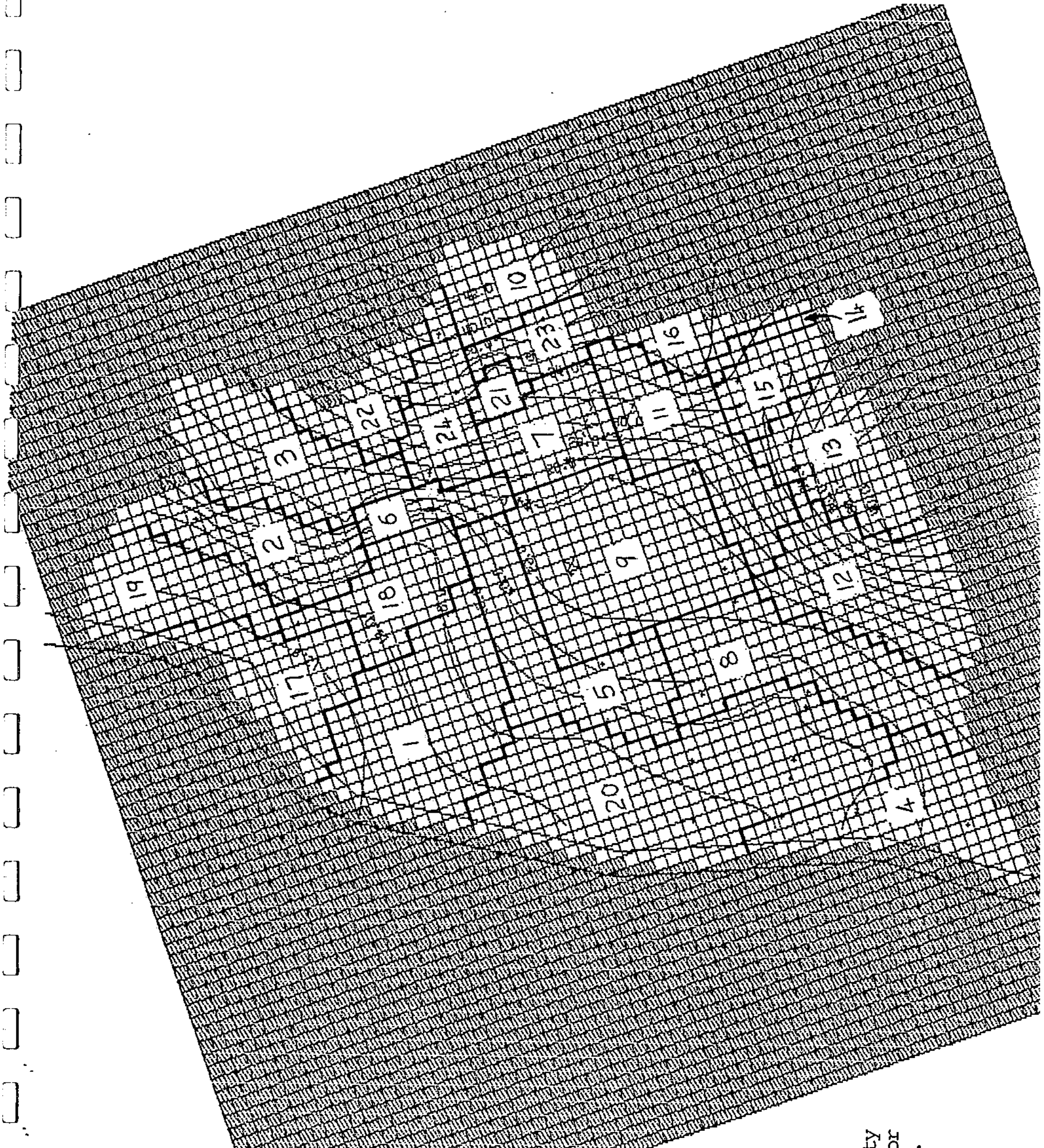

Fig. 3. Transmissivity zones; see Table 1 for transmissivity values.

TABLE 1

Transmissivity Estimates

Zone No. (see Fig. 3)	Transmissivity (m^2/day)	Standard Deviation of ln T
1	45	0.90
2	27	0.13
3	178	0.58
4	13	0.23
5	159	0.19
6	3.6	1.10
7	175	0.08
8	69	0.15
9	844	0.21
10	109	0.44
11	61	0.36
12	12	0.09
13	19	0.13
14	0.10	high
15	15500	high
16	62400	high
17	525	0.20
18	7500	1.14
19	147	0.25
20	371	0.08
21	6650	5.1
22	40	0.75
23	65	0.14
24	8.2	1.55

Note: For transmissivities whose natural logarithms have a standard deviation described as "high", the value output by GENINV far exceeds the range of the linearity assumption on which its calculation is based.

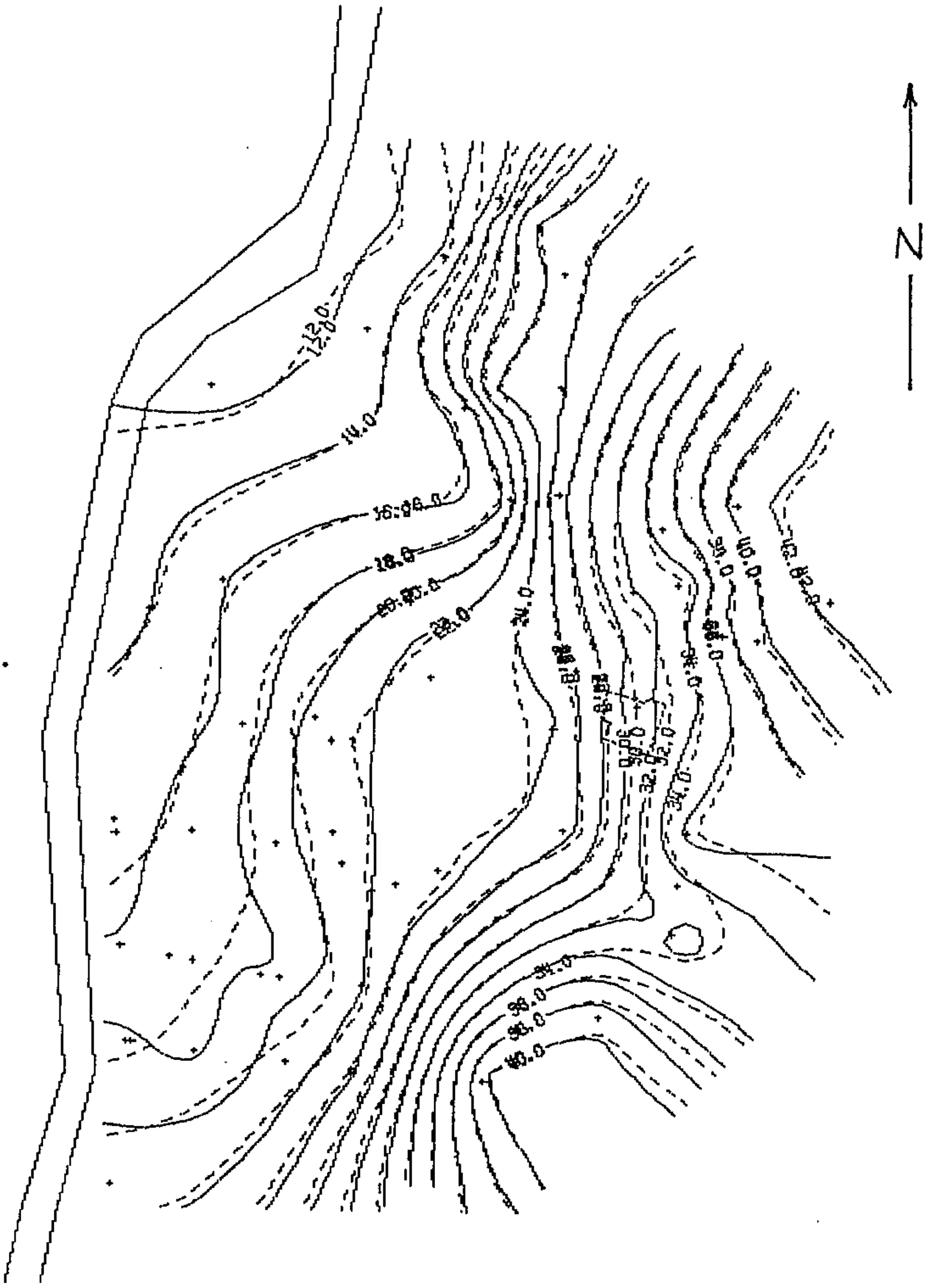

Fig. 4. Measured (full lines) and model (dashed lines) water level contours over Leichhardt Downs. The scale is 1:100000.

The standard deviations for some of the parameters of Table 1 are high. This is to be expected for a model such as that shown in Fig. 3 where a large number of zones are used. High standard deviations will mostly be accompanied by high parameter correlation indicating that two or more parameters could be varied in sympathy with little effect on the contours. Using fewer zones may have resulted in less estimation error for each zone, but worse overall fit between model and field contours.

The cpu time required to achieve the results of Table 1 was 20 hours on a CCI Power 6/32MP computer. This large amount of time may cause problems for users who have more restricted and costly access to cpu time on large mainframe computers. The reason for the large times lies in the fact that the derivatives of head with respect to parameters were calculated by finite differences for the 24 different parameters. This is exacerbated by the fact that towards the end of the inversion process MODFLOW convergence using the SOR solution process was extremely slow, as it appears to be troubled by large transmissivity contrasts. Also, of course, is the fact that there are over 1800 active cells in Fig. 2, an excessive number for the available borehole information. However only 11 iterations of GENINV were required to achieve minimization of the objective function in spite of the fact that the initial value for all transmissivities was taken as $200m^2$/day. It is noteworthy that cpu times as little as 70 minutes have been needed in other cases for grids with over 600 cells and up to 9 transmissivity zones.

CONCLUSIONS

Wherever MODFLOW is used for modelling groundwater flow, GENINV can be used for parameter estimation. The benefits of automatic parameter estimation are:

(a) it requires only that parameter boundaries be chosen by the user, not the parameters themselves; hence a huge saving in time is achieved over tedious and uncertain manual calibration;

(b) the user can be certain that, given his parameter boundaries and prior estimates, the best possible fit between field and model data has been achieved; unrealistic parameter estimates indicate that either field or non-estimated data is erroneous or that parameter boundaries are incorrect;

(c) parameter stochastic properties are provided so that parameter uncertainties and interdependencies can be established.

Once parameters have been estimated using GENINV, they can be used immediately by MODFLOW, this compatibility resulting from the fact that MODFLOW is the forward modelling processor for GENINV.

It is anticipated that GENINV will be of great use in groundwater modelling in the BRIA, whether on a regional or farm scale. The arduous task of model calibration prior to use has been greatly facilitated. As well as this the modeller is in a better position to include estimates of uncertainty in predicting the effects of irrigation, as well as to update model parameters and predictions as more information comes to hand.

REFERENCES

Bard, Y., 1974. Nonlinear Parameter Estimation. Academic Press, New York.

Carrera, J. and Neuman, S.P., 1986. Estimation of Aquifer Parameters under Transient and Steady State Conditions: 1. Maximum Likelihood Method Incorporating Prior Information. Water Resources Research, Vol. 22, No. 2, pp. 199-210.

Carrera, J. and Neuman, S.P., 1986. Estimation of Aquifer Parameter under Transient and Steady State Conditions: 2. Uniqueness, Stability and Solution Algorithms. Water Resources Research, Vol. 22, No. 2, pp. 211-227.

Carrera, J. and Neuman, S.P., 1986. Estimation of Aquifer Parameters under Transient and Steady State Conditions: 3. Application to Synthetic and Field Data. Water Resources Research, Vol. 22, No. 2, pp. 228-242.

Doherty, J., 1989. Leichhardt Downs Electromagnetic Survey. ACTFR Report No. 89/01.

Shaw, R.J. and Thorburn, P.J., 1985. Prediction of Leaching Fraction from Soil Properties, Irrigation Water and Rainfall. Irrigation Science, Vol. 6, pp. 73-83.